SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Mammalian Cell Culture and Transfection

P19 cells were grown in MEMα (Invitrogen), 10% fetal calf serum, and 1X Penicillin/Streptomycin (Invitrogen). These cells were transiently transfected using Lipofectamine 2000 (Invitrogen), following the manufacturer's instructions. HeLa and mEFs were cultured in DMEM (Invitrogen), 10% fetal calf serum, and 1X Penicillin/Streptomycin (Invitrogen). Primary mNSCs were isolated from E14.5 and cultured as described (Lou et al., 2014). Cells were cultured at 37°C with 5% CO_2. P19 and HeLa cells were transfected using Lipofectamine 2000 (Invitrogen) and TransiT-2020 (Mirus), respectively, according to manufacturer's protocol. mEFs and mNSCs were electroporated using Nucleofector Kit for Mouse Neural Stem Cells (Lonza) according to manufacturer's protocol. shUpf3a and shControls were ordered from Thermofisher, targeting from nucleotide 396 of the Upf3a mRNA. siUpf3a and siControls were ordered from GE Dharmacon (ON-TARGETplus SMART-pool siRNA). Stable shUPF3A knockdown HeLa cells described in Chan et al., 2007.

Generation of *Upf3a*-Mutant Mice

Upf3a exon 3 (114 bp) was selected to be flanked by *loxP* sites, with a 2.9 kb fragment upstream and a 4.8 kb fragment downstream of this exon used as homologous arms for embryonic cell (ESC) gene targeting. The targeting vector included a Neomycin resistance cassette flanked by a *FRT* site, which upon recombination by CRE generated a truncated from of UPF3A protein lacking both

the UPF2 and EJC-interacting domains (Kadlec et al., 2004). ESC gene targeting was performed in a 129 ES cell line at the UCSD Transgenic Core. Southern blot analysis was performed with 2 probes outside of the homologous recombination arms to screen for successfully recombined ESC clones (Figure S5A). Two selected ESC clones were individually injected into mouse blastocyst by the UCSD Transgenic Core. Male chimeras with the highest fraction of agouti fur from both parent ES cell clones were mated to *C57BL/6* females (Jackson Laboratory) and germline transmission was determined in the F1 generation by the presence of agouti fur color and a positive band corresponding to the mutant allele by PCR. To generate complete knockout mice, the F1 heterozygote mice were mated to animals from the mouse line *B6.FVB-Tg(EIIa-Cre)C5379mgd/J* (Jackson Laboratory), which express Cre recombinase at an early stage of embryogenesis (Lakso et al., 1996). The F2 mice were back-crossed with *C57BL/6* mice to eliminate the *Cre* allele, and then the F3 mice were inbred to obtain homozygotes.

To generate germ cell- and olfactory epithelium-specific conditional knockout mice, the F1 mice were mated with the mouse line *129S4/SvJaeSor-Gt(ROSA)26Sor*$^{tm1(FLP1)Dym}$*/J* (Jackson Laboratory), which ubiquitously express flippase. The F2 mice were back-crossed to *C57BL/6* mice to eliminate the *flp* allele, and the F3 mice were mated with *Tg(Stra8-cre)1Reb/J* (Jackson Laboratory) or *FVB.Cg-Tg(KRT5-cre/ERT2)2Ipc/JeldJ* (Jackson Laboratory), and then the F4 mice were inbred to obtain various genotypes, including *Upf3a*-cKO mice, as described in the text. *Krt5-Cre*$^{+/-}$*;Upf3a*$^{fl/fl}$ mice were fed a 2-week

tamoxifen diet to induce Cre expression (Harlan). Primers used to generate Southern blot probes and to genotype *Upf3a*-mutants are listed in Supplemental Table 4.

RNA, Luciferase, and Protein Analysis

Total cellular RNA was isolated from cells and tissues using Trizol (Invitrogen), as described (Lou et al., 2014). RNA samples were treated with DNaseI to remove genomic contaminants (Ambion). RNA obtained from germ cell subsets were isolated as previously described (Wang et al., 2005). qPCR analysis was done in triplicate using iScript reverse transcriptase (Bio-Rad) and SYBR-green Real-Time PCR kits (Bio-Rad), as described (Lou et al., 2014). Unless otherwise noted, all qPCR experiments were normalized to the level of *Rpl19* RNA. RT-PCR was performed using Taq Polymerase (Denville).

NMD activity was measured using the NMD reporter plasmids pCI-Neo-WT PTC (-) and pCI-Neo-NS39 PTC (+), which both express Renilla luciferase (Boelz et al., 2006). They were cotransfected with pCI-Neo-FLY, a Firefly luciferase control plasmid. Luciferase level was measured 24 hours after transfection using the Dual Luciferase Reporter Assay System (Promega). Protein tethering NMD activity was measured using MS-HRNP-A1, MS-UPF1, MS-UPF2 expression constructs cotransfected with a β-globin control vector lacking MS-binding sites, and then analyzed by Northern blotting as previously described (Lykke-Andersen et al., 2000). To determine the half-life of

endogenous and NMD reporter mRNAs, P19 cells were treated with actinomycin D (5omyci; Sigma) 24 hours after transfection.

Western blot analysis was performed as previously described (Chan et al., 2007). Histology and immunofluorescence of testis sections were performed as previously described (Song et al., 2012). To assay for apoptotic cells *in vivo,* the Apoptag kit was used in accordance with the manufacturer's instructions (Millipore/Chemicon). Immunofluorescence staining of mouse olfactory epithelium was performed as described (Stephan et al., 2012). Isolation of mouse early mouse embryos and embryo immunofluorescence were performed as described previously (Chavez et al., 2014). Brightfield and epifluorescence microscopy were performed with a Leica AF6000 microscope using Leica software. Confocal microsocopy was performed using an Olympus Gemini FlowView microscope at the National Center for Microscopy and Imaging Research (UCSD). All primer sequences are provided in Supplemental Table 4, antibodies are provided in Supplemental Table 5, and expression vectors were obtained from the Andreas Kulozik Laboratory (Kunz et al., 2006).

RNA-Seq Library Construction and Data Analysis

RNA-seq analysis was performed on RNA isolated from P19 cells transfected with siControl or siUPF3A (GE Dharmacon). 24 h after transfection, cells were treated with actinomycin D and RNA was isolated 0, 15, 105, and 225 minutes post-actinomycin D treatment. *Upf3a* was depleted by 70% according to steady state real time PCR results. RNA was isolated with Trizol and total RNA quality

was assessed using an Agilent Bioanalyzer. Samples with an RNA Integrity Number (RIN) of 8 or greater were used to generate RNA libraries using Illumina's TruSeq RNA Sample Preparation Kit, following the manufacturer's specifications (RNA fragmentation time adjusted to 5 minutes). RNA libraries were multiplexed and sequenced at a concentration of 10 pM with 50 base pair (bp) single end reads to a depth of approximately 50 million reads per sample on an Illumina HiSeq2500 machine. Raw data analysis was performed using the Tuxedo suite, encompassing the Bowtie2, Tophat2, and Cufflinks programs, and using RStudio for subsequent procedures (Trapnell et al., 2010). Approximately 20 million reads were generated per sample, and mapped to the Refseq mm10 genome. Reads were filtered, such that genes without at least one FPKM were removed from the analysis, and overlapping RefSeq transcripts were collapsed giving one expression value per gene locus. The count data was normalized and differential expression was performed using Cufflinks. Genes with an adjusted P value less than 0.05 were considered differentially expressed unless otherwise noted. All functional enrichment analyses were generated using Metascape (http://metascape.org) (Tripathi et al., 2015). Logarithmic regression analysis of normalized count data gave corresponding slopes and y-intercepts for siUpf3a and control datasets. Highly correlative ($R^2 > 0.7$) genes with negative slopes for both control and siUPF3A conditions were defined as identified. Destabilized genes were defined as having a slope change over 10%, with control slope greater than siUPF3A, whereas stabilized genes were defined as having a slope change over 10%, with siUPF3A slope greater than control. Unchanged genes

had a slope change less than 10%. For analysis of NMD-inducing features, corresponding Refseq transcripts were converted into Ensembl transcript IDs and sequences were obtained using the UCSC Table Browser. The criteria that were used to identify transcripts with NMD-inducing features from the Ensembl database were previously described (Lou et al., 2014); the only modifications were that uORFs at least 16 amino acids in length and 3' UTR at least 1000 nt in length were considered as NMD-inducing features.

For pachytene-enriched transcripts, Affymetrix IDs from GSE4193 were converted to Ensembl gene IDs and then Ensembl transcript IDs using the biomaRt package in R's programming environment. During the conversion, only transcripts with a 'protein_coding' biotype were selected. Once the complete repertoire of protein coding transcripts was obtained, UCSC's Table Browser identified transcript sequences. NMD features were identified as described above.

Phylogenetic Tree Construction

Phylogenetic analysis of UPF3 protein-coding sequences was done using 17 representative animal taxa. Vertebrates have two copies (i.e., paralogs) of this gene, Upf3a and Upf3b, as compared to the single copy genes present in invertebrates, consistent with the notion that these two paralogs were generated during the genome duplication events early in vertebrate history after the split from invertebrates (Dehal et al., 2002; Smith et al., 2013). Selected UPF3 sequences were downloaded from Genbank, aligned with MUSCLE. Poorly

aligned regions were removed using Gblocks and analyzed with PAUP v. 4.0b10 for Macintosh (Swofford et al., 2002). Distance analysis used minimum evolution as the optimality criterion (heuristic search with tree-bisection-reconnection and random addition sequence with 100 replications), and mean character difference as the distance measure. Bootstrap analysis used 1000 replicates. Values above 90% are indicated at the respective nodes. The tree is rooted on the five protostome taxa.

Coimmunoprecipitation (CoIP) Assay

CoIP was performed as described in Frank et al., 2010. In brief, after transfection, cells were lysed in a hypotonic lysis buffer, followed by a 4 hour incubation with rabbit polyclonal anti-UPF1 antibody (gifted from Jens Lykke-Andersen) conjugated to protein A sepharose beads (Life Technologies, Carlsbad, CA). Lysates were washed 8 times with 0.05% Triton X-100, 50 mM Tris-HCl pH 7.5, 150 mM NaCl. Lysates were resuspended in 20 µl of load buffer (100 mM Tris-HCl (pH 6.8), 4% SDS, 0.2% bromophenol blue, 20% glycerol). Samples were analyzed by SDS-PAGE and Western blot using rabbit polyclonal anti-MLN51 antibody (Bruno et al., 2011).

Germ Cell Profiling

Isolation and dissociation of cells from seminiferous tubules were performed as previously described (Bastos et al., 2005; McCarrey et al., 1992). FACS analysis

was performed at the UCSD Human Embryonic Stem Cell Core Facility with a BD Influx Cell Sorter (BD Biosciences).

SUPPLEMENTAL REFERENCES

Dehal, P., Satou, Y., Campbell, R. K., Chapman, J., Degnan, B., De Tomaso, A., … Rokhsar, D. S. (2002). The draft genome of Ciona intestinalis: insights into chordate and vertebrate origins.

Frank, F., Sonenberg, N., & Nagar, B. (2010). Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2.

Lakso, M., Pichel, J.G., Gorman, J.R., Sauer, B., Okamoto, Y., Lee, E., Alt, F.W., and Westphal, H. (1996). Efficient in vivo manipulation of mouse genomic sequences at the zygote stage.

McCarrey, J. R., & Dilworth, D. D. (1992). Expression of Xist in mouse germ cells correlates with X-chromosome inactivation.

Smith, J. J., Kuraku, S., Holt, C., Sauka-Spengler, T., Jiang, N., Campbell, M. S., … Li, W. (2013). Sequencing of the sea lamprey (Petromyzon marinus) genome provides insights into vertebrate evolution.

Swofford et al. (2002). Phylogenetic analysis using parsimony (*and other methods), Version 4.

Trapnell et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation.

Tripathi, S., Pohl, M. O., Zhou, Y., Rodriguez-Frandsen, A., Wang, G., Stein, D. A., … Chanda, S. K. (2015). Meta- and Orthogonal Integration of Influenza "OMICs" Data Defines a Role for UBR4 in Virus Budding.